外显和内隐情绪韵律加工的脑机制：近红外成像研究[1*]

雷震[1]　毕蓉[2]　莫李澄[2]　于文汶[2]　张丹丹[1,2]

([1]西南财经大学中国行为经济与行为金融研究中心，成都 611130, [2]深圳大学心理学院，深圳 518060)

摘　要　准确识别言语中的情绪韵律信息对社会交往非常重要。本研究采用功能近红外成像技术，探索外显和内隐情绪加工条件下愤怒、恐惧、快乐三种情绪韵律加工过程中的大脑皮层神经活动。结果表明，对愤怒、恐惧、快乐韵律进行特异性加工的脑区分别为左侧额极/眶额叶、右侧缘上回、左侧额下回，其中右侧缘上回脑区同时受到情绪和任务的调控。此外，右侧颞中回、颞下回和颞极在情绪外显任务中的激活明显强于内隐任务。本研究的结果部分支持了情绪韵律的层次模型，也对该模型的第三层次，即“额区对语音情绪信息的精细加工需要外显性情绪加工任务参与”提出了质疑。

关键词　情绪；语音韵律；颞上回；眶额叶；额下回；缘上回

1 引言

准确解码情绪信息不但对人类的生存至关重要，而且还能帮助我们在社会交往中感知他人的情绪状态、推测他人的意图。在日常交流中常见的传递情绪的载体包括面孔表情、身体姿态、语音等。对面孔和身体姿势情绪信息的加工通过视觉通道完成，目前其神经机制已经得到了较充分的揭示（参见综述 Calvo & Nummenmaa, 2016; Enea & Iancu, 2016; Hinojosa, Mercado, & Carretié, 2015）。相比于视觉通道，我们对语音中情绪信息的解码机制还不是非常清楚（Liebenthal, Silbersweig, & Stern, 2016）。语音的情绪信息可以通过语义和韵律表达。本研究拟考察情绪韵律加工的神经机制。情绪性韵律（affective prosody）是指说话人不依赖

收稿日期：2020-05-03

* 国家自然科学基金（31970980）、深圳市基础研究自由探索项目（JCYJ20180305124305294）、深港脑科学创新研究院(2019SHIBS0003)、及教育部人文社会科学研究规划基金项目（16XJA790004）支持。

通讯作者：张丹丹, E-mail: zhangdd05@gmail.com。并列第一作者：雷震，毕蓉。

语义和语法结构，借助语音中的音高、音长、音强、重音和语调等多种声学线索的动态变化来表达和传递情绪的方式（Brük, Kreifelts, & Wildgruber, 2011）。语音的情绪韵律线索不但帮助我们理解说话人的语义和情绪状态，当情绪韵律与语义信息冲突时，我们通常依赖前者去推测说话人的意图（Ben-David, Multani, Shakuf, Rudzicz, & van Lieshout, 2016）。例如，当某人用愤怒的情绪韵律说“我真幸运”时，我们会认为说话人想表达愤怒情绪。因此，对语音中情绪韵律的加工与我们的生活和社会交往密切相关（Brük et al., 2011; Früholz, Trost, & Kotz, 2016）。揭示和阐明情绪韵律加工的神经机制，不但帮助我们认识大脑的情绪语音加工过程，也有利于对具有社会功能障碍的患者（自闭症、精神分裂症、抑郁症等）进行早期诊断和康复评估（Knight & Baune, 2019; Lin, Ding, & Zhang, 2018）。此外，揭示大脑对情绪韵律的解码规则是人工合成情绪性语音语调的基础，这将极大促进人工智能和人-机交互界面的发展（Mitchell & Xu, 2015）。

早期关于情绪韵律加工的神经机制模型（Ross, 1981）认为，情绪韵律完全由右半脑完成，其中右侧颞上皮层（superior temporal cortex, STC[2]）中与左脑 Wernicke 区对应的部位负责情绪韵律的感知和理解，右背外侧前额叶中与左脑 Broca 区对应的部位负责产生情绪性语音韵律。随着研究的深入，学者们逐渐发现大脑皮层的其他部分和皮下结构也参与了情绪韵律的加工，逐渐形成了情绪韵律加工的层次模型（hierarchical model）（Brük et al., 2011; Ethofer, et al. 2006; Schirmer & Kotz, 2006; Witteman, Van Heuven, & Schiller, 2012）。该模型认为大脑对情绪韵律的加工分为三个层次：1）对声音敏感的 STC 的中部（mid-STC）接收初级听觉皮层的神经信号，负责提取语音中的声学参数；2）右侧 STC 的后部（p-STC）通过多通道整合的方式识别语音中的情绪信息；3）以双侧额上回（inferior frontal gyrus, IFG）和眶额皮层（orbitofrontal cortex, OFC）为代表的额区负责对语音中的情绪信息进行评估和精细加工。情绪韵律加工的层次模型认为，听觉信息从初级听皮层传至 mid-STC 的过程为自底向上的刺激驱动加工，而在 p-STC 和额区进行的语音信息整合和情绪精细加工需要依赖于注意和外显性情绪评估（Brük et al., 2011）。进一步的，Früholz 等（2016）在对情绪声音（包括非人类发出的声音、人类的非言语声音、韵律语音、音乐）的加工脑机制研究结果进行元分析后指出，情绪声音加工依赖于广泛分布的神经网络，其核心脑区包括杏仁核（Früholz, Hofstetter, Cristinzio, Saj, Seeck, & Vuilleumier, 2015）、初级和次级听觉皮层、右侧 STC（Früholz & Grandjean, 2013a）、IFG（Früholz & Grandjean, 2013b）、OFC（Kotz, Kalberlah,

[2] STC 由颞上回（superior temporal gyrus, STG）、颞中回（middle temporal gyrus, MTG）和颞上沟（superior temporal sulcus, STS）三部分组成。

Bahlmann, Friederici, & Haynes, 2013)、脑岛(Mothes-Lasch, Mentzel, Miltner, & Straube, 2011)。

我们认为在情绪韵律加工的脑机制方面，目前有三个问题需要解决。第一，不同情绪种类的语音韵律的大脑表征。已知人类对愤怒、厌恶、恐惧、快乐和悲伤等基本情绪的加工依赖不完全重叠的脑区，即大脑对不同种类情绪的加工具有相对特异性（Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012），但这一结论主要是依据视觉通道的研究证据得出的（例如面孔表情、身体姿势、情绪词加工等任务）。在情绪的听觉加工领域，特别是情绪韵律加工方面，目前已发现有两项相关研究，他们利用机器学习算法，采用数据驱动的方式对被试在收听不同种类情绪韵律的功能性核磁共振成像(functional magnetic resonance imaging, fMRI）激活模式进行分类（Ethofer, Van De Ville, Scherer, & Vuilleumier, 2009a; Kotz et al., 2013)。结果发现，对多分类（区分不同种类的情绪）有贡献的脑区包括：颞上回（superior temporal gyrus, STG)、颞上沟(superior temporal sulcus, STS)、颞中回(middle temporal gyrus, MTG)、IFG、额中回（middle frontal gyrus, MFG)、前脑岛。这两项研究表明，上述脑区对不同种类情绪韵律的加工具有一定的区分度（即特异性)。另外 Kotz 等人（2013）还对比了多种情绪韵律与中性韵律激活的脑区，发现每种情绪都激活了不同的脑区，例如，愤怒韵律比中性韵律更强地激活了双侧 MFG，快乐韵律比中性韵律更强地激活了左侧 IFG。然而，除了上述三项研究，目前在情绪韵律加工方面已有的研究，大多只关注了某一种情绪韵律与中性韵律的差异性大脑表征，很少考察大脑对不同情绪种类韵律的特异性加工（Frühholz & Grandjean, 2013a)。我们认为，仅比较某种情绪与中性条件的差异性脑区，并不能获得情绪韵律加工的“特异性”脑区。本研究将考察愤怒、恐惧、快乐这三类辨识度较高（Liu & Pell, 2012）且具有代表性的情绪韵律的特异性加工脑区。具体的，我们将对三种情绪韵律条件进行两两比较，尝试找寻仅参与单种情绪韵律加工的脑区。依据相关文献(Frühholz & Grandjean, 2013b)，我们预期在 IFG 等额区可能发现能区分三种情绪韵律的特异性脑区。

第二，“情绪相关”任务与“情绪无关”任务（task relevant/irrelevant）激活脑区的异同。虽然已有一些研究采用外显和内隐任务比较了两种条件下情绪韵律加工参与脑区的差异，但目前矛盾的结果还很多。例如，有研究发现内隐任务激活了 p-STG，而外显任务反而激活了 mid-STG（Frühholz, Ceravolo, & Grandjean, 2012），这与前文所述的层次模型的预期相反。此外，不少研究发现杏仁核、脑岛等情绪脑区在外显任务中的激活比内隐任务更强（Frühholz et al., 2012）；而另一些研究则同时在外显和内隐任务中观察到了情绪主效应，即愤怒等情绪韵律比中性韵律在这些脑区均诱发了更强的神经活动（Bach et al., 2008; Ethofer et al.,

2009b; Quadfieg, Mohr, Mentzel, Miltner, & Straube, 2008）。更重要的是，已有研究在额区的发现也不一致：有研究报道外显和内隐情绪任务均会激活双侧或单侧IFG（Früholz et al., 2012; Steber, König, Stephan, & Rossi, 2020），但也有研究观察到了OFG（Ethofer et al., 2009b; Quadfieg et al., 2008）和IFG（Bach et al., 2008; Beaucousin et al., 2011）等脑区在外显任务中有更强的激活。我们认为，现有的矛盾首先可能是由于任务设置的不合理造成的。上述研究几乎均采用“情绪辨别”作为外显任务，“性别辨别”作为内隐任务。相比于情绪辨别，辨别说话人的性别更简单，因此被试在完成性别辨别任务时可能还不自主的进行了情绪加工、语义加工等，从而在结果中引入了混淆变量。本研究将“性别辨别”替换为“身份辨别”，即语音材料均为女性声音，被试的任务为辨别不同身份的女性，加大了内隐任务的难度，降低了被试进行情绪加工的可能性。其次，已有研究绝大多数采用词语作为刺激材料（例如Bach et al., 2008; Ethofer et al., 2009b; Früholz et al., 2012; Quadfieg et al., 2008; Steber et al., 2020）。考虑到情绪韵律为语音的长时程特征，而单个词的呈现时间非常有限（550~750 ms，多为双音节词），我们认为在情绪韵律的研究中应该采用更长的语音单元（例如句子）作为实验材料。此外，已有的研究有的使用有语义的材料（Beaucousin et al., 2011; Ethofer et al., 2009b; Mitchell, 2007; Quadfieg et al., 2008），有的使用无语义的材料（Bach et al., 2008; Früholz et al., 2012; Steber et al., 2020），这也可能造成对实验结果的干扰，因此在本研究中我们使用了无语义的“伪句”。在提高了内隐任务的认知负荷、延长了情绪韵律材料的呈现时间、排除了语义加工的干扰后，本研究希望能更准确地揭示外显和内隐情绪韵律加工的脑机制差异。由于已有文献存在大量矛盾的结果，我们对此问题无法提出具体的假设（探索性问题）。

第三，基于无噪声脑成像技术的实验证据需进一步累积。已有的绝大部分情绪韵律加工脑机制研究使用了fMRI技术（Früholz et al., 2016）。fMRI是目前脑影像研究中的主流技术，能对全脑的神经活动进行较准确的空间定位。但fMRI在扫描过程中由于梯度线圈定位等环节会产生高强度的噪声，这对情绪韵律加工等需要关注各种声学参数细节的任务会造成一定的影响（Dieler, Tupak, & Fallgatter, 2012）。因此我们建议基于声学感知通道的fMRI研究结果最好能通过其他脑成像技术加以验证。功能近红外光谱成像技术（functional near-infrared spectroscopy, fNIRS）具有无噪声、对头动相对不敏感等特点，很适合用于语音加工研究。本研究希望利用fNIRS技术得到与已有的fMRI研究较一致的情绪韵律加工脑区。

2 方法

2.1 被试

采用 G*Power3.1.7 软件预估样本量，参考已有类似文献的效应量（partial η^2 > 0.10）（Bach et al., 2008; Früholz et al., 2012），计算得到 16 名被试即可达到 0.95 的统计检验力（power）。同时，参考课题组前一项研究的样本量（n = 22）（Zhang, Zhou, & Yuan, 2018），本研究招募了 25 名（其中 13 名女生）大学在校学生（年龄：20.42 ± 2.13 y）参与实验。所有被试均听力正常、右利手。被试排除标准：1）有精神病史；2）有头部创伤史或癫痫病史；3）有严重躯体疾病；4）酒精或药物依赖。实验前告知被试实验仪器及实验内容的相关信息，被试签署了书面知情同意书，实验结束后支付被试一定的报酬。实验方案经深圳大学伦理委员会批准。

2.2 实验材料和实验过程

情绪韵律材料选自中国语音情绪库（Liu & Pell, 2012）。该库的情绪韵律语音为符合汉语语法规则但无意义的“伪句”：每个句子符合主、谓、宾结构，但主谓宾成分均采用无意义的“伪词”，从而消除了句意。每个句子时长约 1~2 s。本研究选取愤怒、恐惧、快乐、中性韵律的材料（每种情绪条件含 30 个伪句）。每种情绪分别由 5~9 个句子合成 10 s 整的语音段。每个 10 s 的情绪语音片段分别由 5~8 句情绪语音（为同种情绪）和 1~4 句中性语音（填充刺激）组成，伪句之间无时间间隔。用 120 个伪句制作愤怒、恐惧和快乐韵律各 5 段 10 s 材料，伪句不重复使用。每个 10 s 的语音片段均含两名女性发声的句子，两个身份对应的句子数量在 15 段材料间不等。

本实验为 3（情绪：恐惧、愤怒、高兴）× 2（任务：外显、内隐）被试内设计。实验中语音刺激通过两个音箱播放（EDIFIER-R26T，中国东莞），音响置于距离被试左右耳前方约 50 cm 处，语音材料播放声强为 60 至 70 dB，平均背景噪声值（无语音材料呈现时）为 30 dB。

实验分为两个任务，任务间的顺序在被试之间进行平衡。“情绪外显加工任务”要求被试对每段 10 s 语音内包含的情绪性句子和中性句子进行计数，每段语音结束后按键作答：“情绪性句子比中性句子的数量多 3 个或以上吗？”（答案“是”、“否”各占 50%）。该任务包含恐惧、愤怒、快乐 3 个 block，block 的顺序在被试间随机排列。每个 block 包含 10 段 10 s 的语音（每个语音材料播放两遍），语音材料以随机顺序呈现，每两段语音材料间的静音时

长为 15 s 减去反应时。“情绪内隐加工任务”要求被试对每段 10 s 语音中的两名女性身份发音的句子进行计数，每段语音播放后按键作答：“两名女性播放的句子的数量相差大于 2 或以上吗？”（答案“是”、“否”各占 50%）。该任务同样包含 3 个 block，每个 block 含 10 段 10 s 的语音。在两个任务中，每个 block 持续约 4 min，block 间被试自主掌握休息时间（可选范围：2~10 min）。为控制语音和语义等无关变量引入的混淆，两个任务使用的 30 段 10 s 语音材料完全相同。

2.3 近红外数据记录

使用多通道 fNIRS 系统以连续波形式记录大脑活动（NirScan, HuiChuang, China）。光学探头的定位采用国际 10/20 系统的 NIRS-EEG 兼容帽（EASYCAP, Herrsching, Germany）。根据已有文献，本研究观测的脑区为双侧额叶和颞叶。我们使用了 13 个发射器和 15 个探测器组成了 37 个有效观测通道（图 1），发射器和探测器之间的平均距离为 3.2 cm（范围为 2.8~3.6 cm）。在记录过程中没有发生探测器信号过饱和的现象。我们定义每个通道的中间点为该通道探测的主要脑区，并以此点为圆心进行每个通道的脑区标定。首先采用 NFRI 工具包（http://brain.job.affrc.go.jp/tools/）计算每个通道的 MNI 坐标，然后在成人 Brodmann Talairach 脑模中（Lancaster et al., 2000）查找对应的脑区。各通道的坐标及脑区标定信息见表 1。

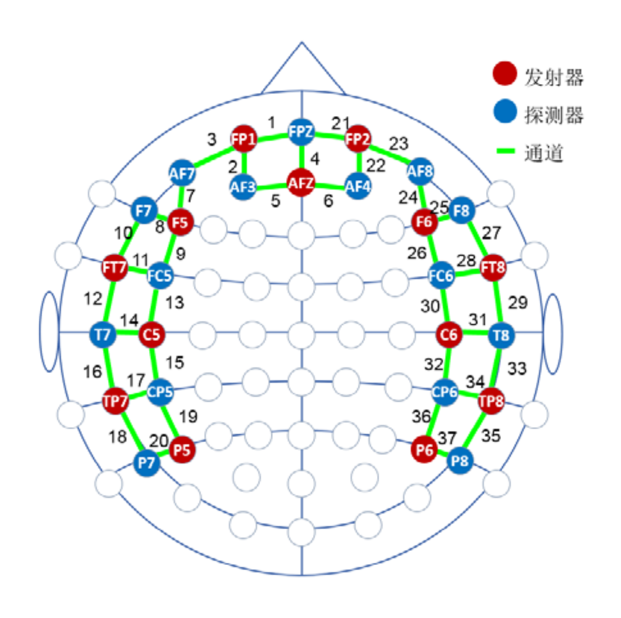


图 1　NIRS 通道排布图

表 1 实验中 37 个 NIRS 通道的空间配准信息

通道	发射器-探测器	MNI 坐标			Brodmann 分区及脑区重合度*
		x	y	z	
1	Fp1-Fpz	-10	68	-5	10 - Frontopolar area (0.62)
2	Fp1-AF3	-25	66	4	10 - Frontopolar area (1.00)
3	Fp1-AF7	-32	62	-8	10 - Frontopolar area (0.58) 11 - Orbitofrontal area (0.42)
4	AFz-Fpz	3	66	11	10 - Frontopolar area (1.00)
5	AFz-AF3	-12	65	20	10 - Frontopolar area (1.00)
6	AFz-AF4	16	65	20	10 - Frontopolar area (1.00)
7	F5-AF7	-46	48	0	10 - Frontopolar area (0.46) 47 - Inferior prefrontal gyrus (0.34)
8	F5-F7	-52	39	0	47 - Inferior prefrontal gyrus (0.62)
9	F5-FC5	-56	27	16	45 - pars triangularis, part of Broca's area (0.64)
10	FT7-F7	-57	21	-13	38 - Temporopolar area (0.68)
11	FT7-FC5	-61	8	2	22 - Superior temporal gyrus (0.61)
12	FT7-T7	-66	-7	-14	21 - Middle temporal gyrus (1.00)
13	C5-FC5	-64	-2	24	6 - Pre-motor and supplementary motor cortex (0.67)
14	C5-T7	-68	-17	8	42 - Primary and auditory association cortex (0.51)
15	C5-CP5	-66	-30	28	40 - Supramarginal gyrus, part of Wernicke's area (0.73)
16	TP7-T7	-69	-31	-9	21 - Middle temporal gyrus (1.00)
17	TP7-CP5	-67	-44	11	22 - Superior temporal gyrus (0.92)
18	TP7-P7	-64	-55	-4	21 - Middle temporal gyrus (0.58) 37 - Fusiform gyrus (0.42)
19	P5-CP5	-60	-56	28	40 - Supramarginal gyrus, part of Wernicke's area (0.58)
20	P5-P7	-58	-68	13	39 - Angular gyrus, part of Wernicke's area (0.42)
21	Fp2-Fpz	14	68	-5	10 - Frontopolar area (0.66)
22	Fp2-AF4	28	66	4	10 - Frontopolar area (1.00)
23	Fp2-AF8	35	63	-8	10 - Frontopolar area (0.63)
24	F6-AF8	49	48	1	10 - Frontopolar area (0.45)
25	F6-F8	54	39	1	47 - Inferior prefrontal gyrus (0.56)
26	F6-FC6	58	25	16	45 - pars triangularis, part of Broca's area (0.69)
27	FT8-F8	59	21	-12	38 - Temporopolar area (0.62)
28	FT8-FC6	63	7	3	22 - Superior temporal gyrus (0.63)
29	FT8-T8	67	-7	-12	21 - Middle temporal gyrus (1.00)
30	C6-FC6	66	-3	24	6 - Pre-motor and supplementary motor cortex (0.66)
31	C6-T8	70	-17	8	42 - Primary and auditory association cortex (0.50)
32	C6-CP6	67	-30	28	40 - Supramarginal gyrus, part of Wernicke's area (0.78)
33	TP8-T8	70	-30	-9	21 - Middle temporal gyrus (0.98)
34	TP8-CP6	68	-43	11	22 - Superior temporal gyrus (0.92)

35	TP8-P8	64	-54	-4	37 - Fusiform gyrus (0.54) 21 - Middle temporal gyrus (0.46)
36	P6-CP6	61	-56	28	40 - Supramarginal gyrus, part of Wernicke's area (0.61)
37	P6-P8	57	-67	13	39 - Angular gyrus, part of Wernicke's area (0.54)

* 部分近红外观测通道可能覆盖了多个脑区，为节省论文空间，本表仅列出重合度高于 0.4 的脑区。

2.4 近红外数据分析

采用 NirSpark 软件（HuiChuang, China）进行数据预处理。1）使用基于小波的运动去伪迹方法（Molavi & Dumont, 2012）对原始光密度数据进行运动伪迹校正；2）使用 0.01~0.20 Hz 的滤波器对数据进行带通滤波；3）基于修正的 Beers-Lambert 定律将滤波后的光密度数据转换为 HbO 和 HbR 的浓度变化数据 Δ[HbO]和 Δ[HbR]。在本研究中，Δ[HbO]比 Δ[HbR]信噪比高，对脑血流的变化更敏感（Tong, Hocke, & Frederick, 2011; Zhang, Chen, Hou, & Wu, 2019），故后续统计分析均采用 Δ[HbO]数据。

使用一般线性模型（general linear model, GLM）解算不同条件下任务相关的 β 值，将 β 值作为衡量相应脑区激活的指标。采用 block 设计，以每段语音 10 s 的时间作为 block 的长度，以此为时间线索与 HRF 进行卷积。统计分析采用 SPSS 20.0 软件（IBM, Somers, USA）。显著性水平设定为 0.05。描述性统计量报告为均值±标准差。对每个通道的 β 值进行 3（情绪：恐惧、愤怒、高兴）× 2（任务：外显、内隐）的重复测量方差分析，采用 Greenhouse-Geisser 进行球形矫正，采用 Bonferroni 进行事后多重比较矫正。最后，在通道间采用 FDR 方法对 p 值进行多重比较矫正，进一步降低假阳性率。

3 结果

反应时（628.35 ± 151.11 ms）和正确率（91.42% ± 6.30%）在各实验条件间无显著差异（$F < 1$）。

脑激活数据分析结果如下。首先，我们在通道 3, 9, 32 发现了情绪主效应（表 2，图 2 及图 3A）。事后多重比较发现，左侧的额极和眶额皮层（frontopolar/orbitofrontal area；即额中回和额上回的前部，通道 3）对愤怒语音最敏感（$F(2,48) = 12.51$, $p < 0.001$, partial $\eta^2 = 0.343$; corrected $p = 0.001$；愤怒 vs. 恐惧 $p = 0.007$；愤怒 vs. 快乐 $p = 0.001$）。左侧的额下回三角部（pars triangularis; Broca 区的一部分；通道 9）对快乐语音最敏感（$F(2,48) = 24.26$, $p < 0.001$, partial $\eta^2 = 0.502$; corrected $p < 0.001$；快乐 vs. 愤怒/恐惧 $ps < 0.001$）。右侧缘上回

（supramarginal gyrus, SMG；通道 32）对恐惧语音最敏感（$F(2,48) = 12.53$, $p < 0.001$, partial $\eta^2 = 0.343$; corrected $p = 0.001$；恐惧 vs. 愤怒 $p = 0.001$；恐惧 vs. 快乐 $p = 0.003$）。

其次，我们在通道 27~29 发现了任务主效应（表 3，图 2 及图 3B）。在右侧颞叶的三个通道内，外显比内隐情绪加工任务引起了更强的皮层活动。其中通道 27 对应于颞极（temporopolar; $F(1,24) = 11.63$, $p = 0.002$, partial $\eta^2 = 0.325$; corrected $p = 0.004$），通道 28 对应于颞上回（STG; $F(1,24) = 26.10$, $p < 0.001$, partial $\eta^2 = 0.521$; corrected $p < 0.001$），通道 29 对应于颞中回（MTG; $F(1,24) = 15.81$, $p = 0.001$, partial $\eta^2 = 0.397$; corrected $p = 0.003$）。

最后，我们在通道 32 发现了任务类型和情绪的交互作用（$F(2,48) = 14.24$, $p < 0.001$, partial $\eta^2 = 0.372$; corrected $p = 0.002$；图 2 及图 3C），该通道对应于右侧 SMG。简单效应分析表明，在外显任务中，右侧 SMG 脑区对恐惧（β 值：0.63 ± 0.43）比对愤怒（0.06 ± 0.44）和快乐韵律（0.10 ± 0.34）更敏感（$F(2,23) = 102.27$, $p < 0.001$, partial $\eta^2 = 0.898$），而该情绪效应在内隐任务中不显著（$F < 1$；恐惧/愤怒/快乐 β 值 = 0.10 ± 0.44 / 0.14 ± 0.66 / 0.12 ± 0.55）。

表 2　情绪主效应结果

通道	脑区	*F*	*p**	愤怒 β 值	恐惧 β 值	快乐 β 值
3	L Frontopolar/orbitofrontal area	12.51	.001	0.21 ± 0.20	0.12 ± 0.23	0.06 ± 0.20
9	L pars triangularis/Broca’s area	24.24	< .001	0.10 ± 0.16	0.10 ± 0.15	0.21 ± 0.15
32	R Supramarginal gyrus	12.48	.001	0.10 ± 0.56	0.36 ± 0.51	0.11 ± 0.45

*The *p* was corrected using FDR method across fNIRS channels.

表 3　任务主效应结果

通道	脑区	*F*	*p**	内隐 β 值	外显 β 值
27	R Temporopolar area	11.62	0.004	0.04 ± 0.36	0.32 ± 0.42
28	R Superior temporal gyrus	26.17	< 0.001	0.05 ± 0.45	0.37 ± 0.43
29	R Middle temporal gyrus	15.84	0.003	-0.03 ± 0.49	0.34 ± 0.53

*The *p* was corrected using FDR method across fNIRS channels.

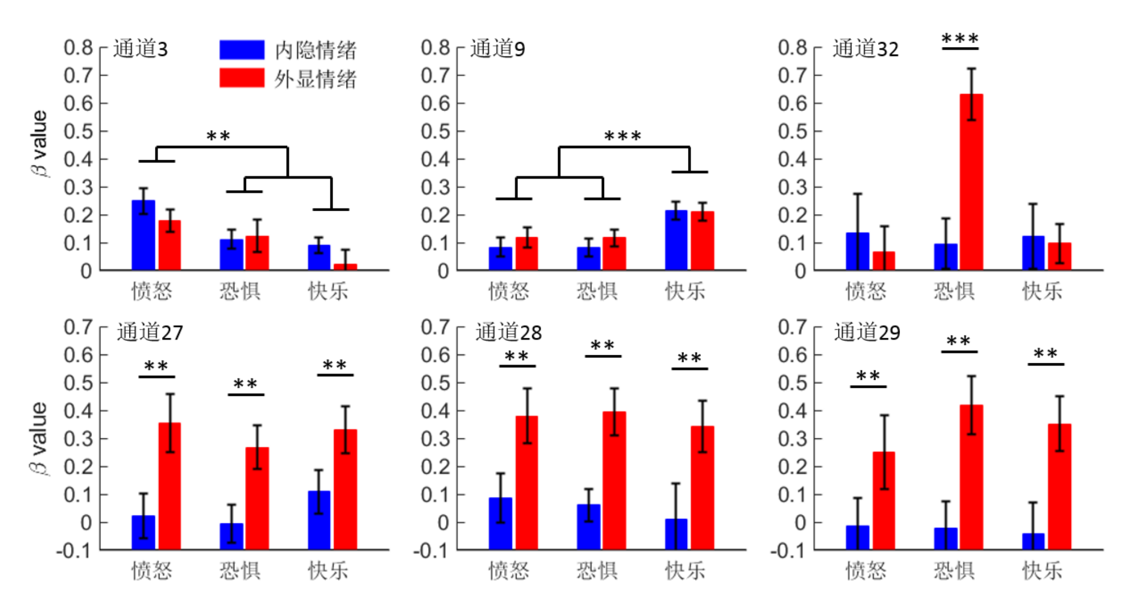


图 2 不同脑区在情绪和任务条件中的激活（仅显示出现显著效应的通道）。图中的 errorbar 表示均值的标准误。

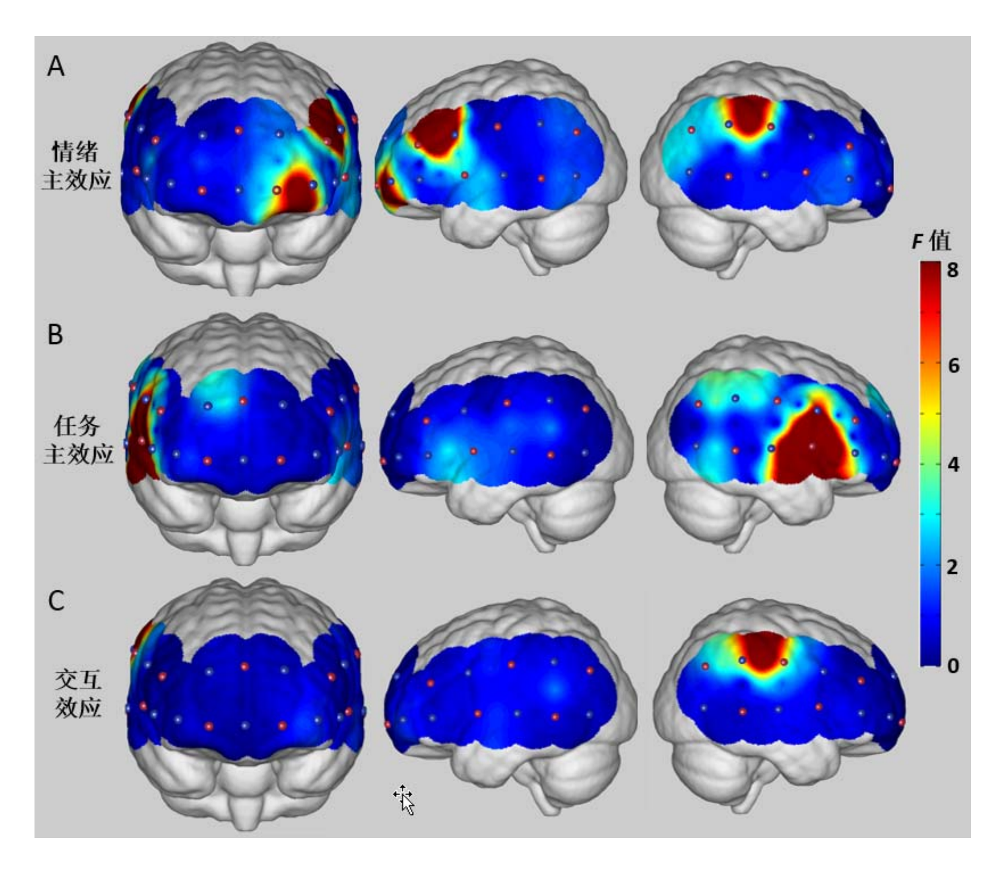


图 3 脑区激活的成像图。

4 讨论

本研究利用 fNIRS 技术，探讨了外显和内隐情绪任务下，情绪韵律加工的大脑皮层神经活动。结果表明，外显情绪任务比内隐任务更强地激活了右侧颞叶的 STG、MTG 和颞极，说明右侧颞叶在情绪任务相关的语音韵律感知中发挥了重要作用。我们还发现，不同种类情绪的辨别性解码依赖于额区的 IFG、OFC，且无论在外显还是内隐情绪任务中这些脑区均会激活。而顶叶的 SMG 脑区仅在外显任务中可辨别出恐惧韵律。

情绪主效应表明，大脑在情绪任务相关和无关条件下均可分辨不同种类的情绪韵律。相较于恐惧和快乐韵律，愤怒韵律更强的激活了左侧的额极和 OFC 脑区。这一发现与此前在情绪韵律辨别任务（Kotz et al., 2013）和被动收听任务（Zhang et al., 2018）中得到的结果非常相似。OFC 不但在愤怒情绪的加工中具有重要作用（Lindquist et al., 2012），同时也负责冲突解决和抑制不恰当的行为（例如攻击行为）（Beyer, Münte, Göttlich, & Krämer, 2015）。OFC 损伤的病人更具有攻击倾向，且在主观情绪状态评估、情绪信息整合、情绪韵律辨识方面存在显著缺陷（Fox et al., 2018; Herpertz et al., 2017; Paulmann, Seifert, & Kotz, 2010）。与愤怒韵律对应 OFC 的结果类似，我们还发现左侧额区的 IFG 可以在外显和内隐任务条件下、从恐惧和愤怒韵律中特异性地区分出快乐韵律，这一结果与此前在成人（Zhang et al., 2018）和新生儿（Zhang et al., 2019）情绪韵律被动收听任务中的发现相符。具体的，我们在通道 9 发现的这一结果，其脑区定位于 IFG 的三角部（pars triangularis），此区域在语义理解和语义-情绪韵律等语言信息整合中具有关键作用（Goucha & Friederici, 2015; Kirby & Robinson, 2017; Schirmer & Kotz, 2006）。一项关注讽刺理解的研究发现，左侧 IFG 在语境、语义和韵律整合加工中发挥了重要作用，当出现赞扬性的语义和负性情绪的韵律时（即讽刺），该脑区激活增加（Matsui et al., 2016）。在另一项听觉情绪和社会判断的研究中，左侧 IFG 被认为同时参与了社会特质评估（信任度评估）和说话者快乐程度评估（Hensel, Bzdok, Müller, Zilles, & Eickhoff, 2015）。Kotz 等人的两项研究发现，与中性韵律相比，快乐韵律激活了左侧 IFG（Kotz et al., 2003; 2013），而 Johnstone 等人的研究则观察到快乐韵律使右侧 IFG 激活增强（Johnstone, Van Reekum, Oakes, & Davidson, 2006）。IFG 激活的左右半球不一致结果可能是实验设置或情绪材料的不同造成的：在 Kotz 等人（2003; 2013）的实验中，被试仅加工语音韵律，而在 Johnstone 等人（2006）的实验中，被试在收听语音韵律的同时还需观察情绪信息一致或不一致的面部表情。本研究的情绪材料与 Kotz 等人（2003; 2013）的研究相似，因此发现快乐韵律特异性的激活了左侧 IFG。由于对情绪条件进行了两两对比

分析，本文发现了愤怒和快乐情绪韵律对应的“特异性脑区”，这一结果在之前那些仅对比某种情绪与中性条件的研究中是未曾得到的。

根据情绪韵律加工的层次模型，在情绪韵律加工的第三层次，OFC、IFG 等额区负责对语音的情绪信息进行精细加工，这一层次需要依赖于注意或外显性情绪评估（Brük et al., 2011）。而本文在 OFC 和 IFG 脑区的发现与此结论不符。回溯已有的采用外显和内隐情绪任务考察情绪韵律加工的研究发现，那些观察到 OFC/IFG 在外显任务中的激活强于内隐任务的研究大多选用了有语义的词（Ethofer et al., 2009b; Quadfieg et al., 2008）或句子（Beaucousin et al., 2011）作为实验材料，而那些观察到额区在外显和内隐任务中均有明显激活的研究往往选用了不具有语义的情绪韵律材料（例如 Früholz et al., 2012; Steber et al., 2020）。因此我们推断语义加工对情绪韵律加工可能造成了干扰或影响。目前我们发现的唯一的一项例外研究为 Bach 等人（2008），他们虽然在实验中采用伪词作为语音材料，但仍然观察到左侧 IFG 在外显比在内隐条件下的激活更强。然而该研究选用性别辨别作为内隐任务，正如引言所述，性别辨别的难度远小于情绪辨别（该研究中，正确率和反应时在两种任务中存在显著差异），这有可能对结果造成影响；此外 Bach 等人（2008）仅招募了 16 名被试进行实验，样本量稍小。本研究使用“身份辨别”作为内隐任务，匹配了内隐和外显条件间的任务难度，而且我们采用无语义的伪句作为实验材料。在此基础上我们发现，额区 OFC/IFG 在外显和内隐任务中均可区分不同种类的情绪韵律，因此本研究能比较有信心地对层次模型第三层次的加工理论提出质疑。

我们的结果还表明，右侧 SMG 脑区对恐惧韵律有相对特异性的加工，这一发现是未曾预期到的。SMG 是参与语言感知和加工的脑区之一，该脑区损伤会导致感觉性失语症。SMG 也参与身体姿势的加工，被认为是镜像神经元系统的一部分（Carlson, 2012）。右侧 SMG 更是与语音韵律加工（Hartwigsen, Baumgaertner, Price, Koehnke, Ulmer, & Siebner, 2010）和情绪感知相关（Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, 2000; Aryani, Hsu, & Jacobs, 2018）。Köchel 等人认为 SMG 与注意、警觉功能密切相关（Köchel, Schöngassner, & Schienle, 2013）。他们的研究采用恐惧、厌恶、中性的非言语声音，发现当被试加工恐惧声音时（例如恐惧或痛苦引发的尖叫），右侧 STG 和双侧 SMG 的激活显著增加。此外，Patel 等人（2018）发现，情绪韵律的表达功能受损与 SMG 受损有关。更值得关注的是，SMG 是本研究中观察到的唯一一个同时受情绪和任务调控的脑区：右侧 SMG 仅在外显情绪任务中表现出对恐惧韵律的特异性加工。一方面，在此前的成人情绪韵律被动收听任务中（Zhang,

Zhou, & Yuan, 2018）我们并未观察到右侧 SMG 对恐惧韵律的敏感性，这也说明该脑区对恐惧情绪的加工需要相当程度的情绪相关任务的参与。另一方面，右侧 SMG 对恐惧韵律的特异性加工与课题组此前在新生儿研究中的发现一致（Zhang, Zhou, Hou, Cui, & Zhou, 2017; Zhang et al., 2019），这说明人类在刚出生时其大脑即可对恐惧情绪韵律进行快速反应，此过程与是否有情绪任务参与无关。结合我们在成人和新生儿实验中对恐惧情绪韵律的发现，我们认为这可能反映了人类在进化过程中发展出的对恐惧情绪的先天性加工偏向，即出生时可以自动地对恐惧情绪产生敏感的脑反应，而随着年龄的增长，对恐惧韵律的加工（而非恐惧面孔的加工）则需要主动的外显情绪任务参与。

本研究的任务主效应的表明，相较于内隐任务，右侧 MTG、STG、颞极在情绪外显任务下激活显著增强，这一结果与情绪韵律加工的层次模型（Bach et al., 2008; Brük et al., 2011; Ethofer, T. et al. 2006; Schirmer & Kotz, 2006; Witteman et al., 2012）完全吻合，也符合情绪语音韵律加工的右侧化理论（见综述 Belyk & Brown, 2014）。右侧 STC 是“情绪语音脑区”（emotional voice area; Ethofer et al., 2012; Liebenthal et al., 2016）的最主要结构，是情绪性语音解码的关键脑区（见综述 Früholz & Grandjean, 2013a）。STC 的低级结构（初级听觉皮层和 mid-STC）负责解析情绪声音（包括语音、人类非语音、自然界声音）的听觉特征，而 STC 的高级结构将这些解析出的音频特征整合并构建出对情绪语音的感知（Früholz et al., 2016; Schirmer & Kotz, 2006）。已有研究表明 STC 对情绪性语音比对中性语音有更强的激活（Bach et al., 2008; Brück et al., 2011; Ethofer et al., 2009b; Früholz et al., 2012; Kotz et al., 2003; Mothes-Lasch et al., 2011; Witteman et al., 2012）。本研究发现，STC 对情绪韵律的外显性加工有明显的右侧化优势，与已有文献一致（Früholz et al., 2016; Kotz et al., 2013），也说明右脑对语音中的慢变信号和超音段特征敏感（Witteman et al., 2012）。考虑到右侧 STG 是经典的音高（pitch）加工脑区（Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002），本文观察到的外显情绪任务效应也可能反应了右侧 STG 在情绪辨别任务中对音高进行主动地辨别加工(愤怒、恐惧、快乐韵律的音高均高于中性韵律；音频材料信息见 Liu & Pell, 2012)。本研究发现的情绪韵律外显性加工的颞区右侧化优势可帮助临床对脑疾病进行诊断，例如判断癫痫的病灶、评估癫痫及大脑器质性病变之后的脑功能可塑性（Alba-Ferrara, Kochen, & Hausmann, 2018）。

综上所述，本研究发现愤怒、恐惧、快乐韵律加工的特异性脑区分别位于左侧 OFC、右侧 SMG 和左侧 IFG。我们还发现右侧 STC（包含 MTG、STG 等）以及右侧 SMG 在外

显性情绪韵律任务中的重要作用。本研究的结果部分支持了情绪韵律加工的层次模型，也对该模型的第三层次，即“额区对语音情绪信息的精细加工需要外显性情绪任务参与”提出了质疑。此外，由于 fNIRS 技术本身的局限性，本研究无法探测到大脑较深部的、与情绪韵律加工密切相关的脑区（例如颞上沟和杏仁核）。我们建议本领域的探索除了依赖 fMRI、fNIRS、脑磁图等多种影像技术，还需结合脑损伤研究、经颅磁刺激技术（特别是最新发展起来的深部经颅磁刺激）等进一步展开。

参考文献

Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *The Journal of Neuroscience, 20*(7), 2683–2690.

Alba-Ferrara, L., Kochen, S., & Hausmann, M. (2018). Emotional prosody processing in epilepsy: Some insights on brain reorganization. *Frontiers in Human Neuroscience, 12*, 92.

Aryani, A., Hsu, C. T., & Jacobs, A. M. (2018). The sound of words evokes affective brain responses. *Brain Sciences, 8*(6), 94.

Bach, D. R., Grandjean, D., Sander, D., Herdener, M., Strik, W. K., & Seifritz, E. (2008). The effect of appraisal level on processing of emotional prosody in meaningless speech. *Neuroimage, 42*(2), 919–927.

Beaucousin, V., Zago, L., Herve, P. Y., Strelnikov, K., Crivello, F., Mazoyer, B., & Tzourio-Mazoyer, N. (2011). Sex-dependent modulation of activity in the neural networks engaged during emotional speech comprehension. *Brain Research, 1390*, 108–117.

Ben-David, B. M., Multani, N., Shakuf, V., Rudzicz, F., & van Lieshout, P. H. (2016). Prosody and semantics are separate but not separable channels in the perception of emotional speech: test for rating of emotions in speech. *Journal of Speech Language and Hearing Research, 59*(1), 72–89.

Beyer, F., Munte, T. F., Gottlich, M., & Kramer, U. M. (2014). Orbitofrontal cortex reactivity to angry facial expression in a social interaction correlates with aggressive behavior. *Cerebral Cortex, 25*(9), 3057–3063.

Belyk, M., & Brown, S. (2014). Perception of affective and linguistic prosody: An ALE meta-analysis of neuroimaging studies. *Social Cognitive and Affective Neuroscience, 9*, 1395–1403.

Brük, C., Kreifelts, B., & Wildgruber, D. (2011). Emotional voices in context: A neurobiological model of multimodal affective information processing. *Physics of Life Reviews, 8*, 383–403.

Calvo, M. G., & Nummenmaa, L. (2016). Perceptual and affective mechanisms in facial expression recognition: An integrative review. *Cognition and Emotion, 30*, 1081–1106.

Dieler, A. C., Tupak, S. V., & Fallgatter, A. J. (2012). Functional near-infrared spectroscopy for the assessment of speech related tasks. *Brain and Language, 121*(2), 90–109.

Enea, V., & Iancu, S. (2016). Processing emotional body expressions: state-of-the-art. *Social Neuroscience, 11*(5), 495–506.

Ethofer, T., Bretscher, J., Gschwind, M., Kreifelts, B., Wildgruber, D., & Vuilleumier, P. (2012). Emotional voice areas: Anatomic location, functional properties, and structural connections revealed by combined fMRI/DTI. *Cerebral Cortex, 22*, 191–200.

Ethofer, T., et al. (2006). Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. *Neuroimage, 30*, 580–587.

Ethofer, T., Van De Ville, D., Scherer, K., & Vuilleumier, P. (2009a). Decoding of emotional information in voice-sensitive cortices. *Current Biology, 19*(12), 1028-1033.

Ethofer, T., Kreifelts, B., Wiethoff, S., Wolf, J., Grodd, W., Vuilleumier, P., & Wildgruber, D. (2009b). Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. *Journal of Cognitive Neuroscience, 21*, 1255–1268.

Fox, K. C. R., Yih, J., Raccah, O., Pendekanti, S. L., Limbach, L. E., Maydan, D. D., & Parvizi, J. (2018). Changes in subjective

experience elicited by direct stimulation of the human orbitofrontal cortex. *Neurology, 91*(16), e1519–e1527.

Früholz, S., Trost, W., & Kotz, S. A. (2016). The sound of emotions - Towards a unifying neural network perspective of affective sound processing. *Neuroscience and Biobehavioral Reviews, 68*, 96–110.

Früholz, S., & Grandjean, D. (2013a). Multiple subregions in superior temporal cortex are differentially sensitive to vocal expressions: a quantitative meta-analysis. *Neuroscience and Biobehavioral Reviews, 37*, 24–35.

Früholz, S., & Grandjean, D. (2013b). Processing of emotional vocalizations in bilateral inferior frontal cortex. *Neuroscience and Biobehavioral Reviews, 37*(10), 2847–2855.

Früholz, S., Hofstetter, C., Cristinzio, C., Saj, A., Seeck, M., & Vuilleumier, P. (2015). Asymmetrical effects of unilateral right or left amygdala damage on auditory cortical processing of vocal emotions. *Proceedings of the National Academy of Sciences of the United States of America, 112*(5), 1583–1588.

Früholz, S., Ceravolo, L., & Grandjean, D. (2012). Specific brain networks during explicit and implicit decoding of emotional prosody. *Cerebral Cortex, 22*, 1107–1117.

Goucha, T., & Friederici, A. D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca's Area. *Neuroimage, 114*, 294–302.

Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010). Phonological decisions require both the left and right supramarginal gyri. *Proceedings of the National Academy of Sciences of the United States of America, 107*(38), 16494–16499.

Hensel, L., Bzdok, D., Müller, V. I., Zilles, K., & Eickhoff, S. B. (2015). Neural correlates of explicit social judgments on vocal stimuli. *Cerebral Cortex, 25*(5), 1152–1162.

Herpertz, S. C., Nagy, K., Ueltzhöffer, K., Schmitt, R., Mancke, F., Schmahl, C., & Bertsch, K. (2017). Brain mechanisms underlying reactive aggression in borderline personality disorder-sex matters. *Biological Psychiatry, 82*(4), 257–266.

Hinojosa, J. A., Mercado, F., & Carretié, L. (2015). N170 sensitivity to facial expression: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 55*, 498–509.

Johnstone, T., Van Reekum, C. M., Oakes, T. R., & Davidson, R. J. (2006). The voice of emotion: an FMRI study of neural responses to angry and happy vocal expressions. *Social Cognitive and Affective Neuroscience, 1*(3), 242–249.

Kirby, L. A. J., & Robinson, J. L. (2017). Affective mapping: An activation likelihood estimation (ALE) meta-analysis. *Brain and Cognition, 118*, 137–148.

Knight, M. J., & Baune, B. T. (2019). Social cognitive abilities predict psychosocial dysfunction in major depressive disorder. *Depression and Anxiety, 36*(1), 54–62.

Kotz, S. A., Kalberlah, C., Bahlmann, J., Friederici, A. D., & Haynes, J. D. (2013). Predicting vocal emotion expressions from the human brain. *Human Brain Mapping, 34*, 1971–1981.

Kotz, S. A., et al. (2003). On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain and Language, 86*, 366–376.

Köchel, A., Schöngassner, F., & Schienle, A. (2013). Cortical activation during auditory elicitation of fear and disgust: a near-infrared spectroscopy (NIRS) study. *Neuroscience Letters, 9*(549), 197–200.

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., … Fox, P. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping, 10*(3), 120–131.

Liebenthal, E., Silbersweig, D. A., & Stern, E. (2016). The Language, Tone and prosody of emotions: neural substrates and dynamics of spoken-word emotion perception. *Frontiers in Aging Neuroscience, 10*, 506.

Lin, Y., Ding, H., & Zhang, Y. (2018). Emotional prosody processing in schizophrenic patients: A selective review and meta-analysis. *Journal of Clinical Medicine, 7*(10), 363.

Liu, P., & Pell, M. D. (2012). Recognizing vocal emotions in Mandarin Chinese: a validated database of Chinese vocal emotional stimuli. *Behavior Research Methods, 44*, 1042–1051.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review.

Behavioral and Brain Sciences, 35, 121–143.

Matsui, T., Nakamura, T., Utsumi, A., Sasaki, A. T., Koike, T., Yoshida, Y., & Harada, T., et al. (2016). The role of prosody and context in sarcasm comprehension: Behavioral and fMRI evidence. *Neuropsychologia, 87*, 74–84.

Mitchell, R. L., & Xu, Y. (2015). What is the value of embedding artificial emotional prosody in human-computer interactions? Implications for theory and design in psychological science. *Frontiers in Psychology, 6*, 1750.

Mitchell, R. L. (2007). fMRI delineation of working memory for emotional prosody in the brain: commonalities with the lexico-semantic emotion network. *Neuroimage, 36*(3), 1015–1025.

Mothes-Lasch, M., Mentzel, H. J., Miltner, W. H. R., & Straube, T. (2011). Visual attention modulates brain activation to angry voices. *Journal of Neuroscience, 31*, 9594–9598.

Ross, E. D. (1981). The aprosodias. Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology, 38*(9), 561–569.

Patel, S., Oishi, K., Wright, A., Sutherland-Foggio, H., Saxena, S., Sheppard, S. M., & Hillis, A. E. (2018). Right hemisphere regions critical for expression of emotion through prosody. *Frontiers in neurology, 9*, 224.

Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron, 36*, 767–776.

Paulmann, S., Seifert, S., & Kotz, S. A. (2010). Orbito-frontal lesions cause impairment during late but not early emotional prosodic processing. *Social Neuroscience, 5*(1), 59–75.

Quadfieg, S., Mohr, A., Mentzel, H. J., Miltner, W. H., & Straube, T. (2008). Modulation of the neural network involved in the processing of anger prosody: the role of task-relevance and social phobia. *Biological Psychology, 78*, 129–137.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10*, 24–30.

Steber, S., König, N., Stephan, F., & Rossi, S. (2020). Uncovering electrophysiological and vascular signatures of implicit emotional prosody. *Scientific Reports, 10*(1), 5807.

Tong, Y., Hocke, L. M., & Frederick, B. deB., (2011). Isolating the sources of widespread physiological fluctuations in functional near-infrared spectroscopy signals. *Journal of Biomedical Optics, 16*(10), 106005.

Witteman, J., Van Heuven, V. J., & Schiller, N. O. (2012). Hearing feelings: a quantitative meta-analysis on the neuroimaging literature of emotional prosody perception. *Neuropsychologia, 50*, 2752–2763.

Zhang, D., Chen, Y., Hou, X., & Wu, Y. J. (2019). Near-infrared spectroscopy reveals neural perception of vocal emotions in human neonates. *Human Brain Mapping, 40*(8), 2434–2448.

Zhang, D., Zhou, Y., Hou, X., Cui, Y., & Zhou, C. (2017). Discrimination of emotional prosodies in human neonates: A pilot fNIRS study. *Neuroscience Letters, 658*, 62–66.

Zhang, D., Zhou, Y., & Yuan, J. (2018). Speech prosodies of different emotional categories activate different brain regions in adult cortex: an fNIRS study. *Scientific Reports, 8*(1), 218.

The brain mechanism of explicit and implicit processing of emotional prosodies：An fNIRS study

LEI Zhen[1]; BI Rong[2]; MO Licheng[2]; YU Wenwen[2]; ZHANG Dandan[1,2]

[1] China Center for Behavioral Economics and Finance & School of Economics, Southwestern University of Finance and Economics, Chengdu 611130, China, [2] College of Psychology, Shenzhen University, Shenzhen 518060, China

Abstract Emotional expressions of others embedded in speech prosodies are important for social interactions. Affective prosody refers to a way to express and convey emotions through the dynamic changes of various acoustic cues such as pitch, intensity, stress, and intonation in speech, without relying on vocabulary and grammatical structure. Previous studies have shown that STC, IFG, OFC, and other cerebral cortex and subcortical structures are involved in emotional prosody processing, and gradually formed a hierarchical model. However, existing studies on the neural mechanism of emotional prosody processing mostly focus on the difference between non-neutral emotional prosody and neutral prosody, while the comparison between various non-neutral emotional prosody is less investigated. Besides, the differences involved in brain regions of emotional prosody processing under explicit and implicit tasks are still not clear. Furthermore, it is necessary to further accumulate experimental evidence based on noise-free brain imaging technology, such as the noise-free features of fNIRS are especially suitable for speech processing research.

This study used functional near-infrared spectroscopy to investigate how speech prosodies of different emotional categories are processed in the cortex under different task conditions. A group of 25 college students participated in this study with a 3 (emotion: anger vs. fearful vs. happy) by 2 (task focus: explicit vs. implicit) within-participant factorial design. We manipulated task focus by adopting two different tasks, with emotional discrimination task as explicit condition and sex discrimination task as implicit condition. Ten phonological materials for each of anger, fearful, and happy prosody were selected from the Chinese Speech Emotion Database and consisted of the corresponding emotional prosodies and neutral prosodies. The emotional explicit task was to count the emotional and neutral sentences contained in each 10-second speech, and the emotional implicit task was to count the sentences played by two women in each 10-second speech. A multi-channel fNIRS system was used to record brain activity in a continuous waveform. According to existing literature, the brain regions observed in this study are the bilateral frontal and temporal lobes. Therefore, we used 13 emitters and 15 detectors to form 37 effective observation channels.

We first adopted NirSpark-2442 software to preprocess the data, and then conducted general linear model analyses to calculate the cortical activation related to the task. The results showed that the brain activation was significantly higher when anger was contrasted to fearful and happy prosody in left frontal pole / orbitofrontal cortex, and when happy was contrasted to fearful and anger prosody in left inferior frontal gyrus, and when fearful was contrasted to anger and happy

prosody in right supramarginal gyrus. Importantly, there was an interaction between emotion and task. In the explicit task, cortex activity in the right supramarginal gyrus was more sensitive to fearful than to anger and happy prosodies. But no similar results were found under anger and happy prosody. In addition, the brain activation in temporopolar, superior temporal gyrus, and middle temporal gyrus with the explicit task was greater than that in the implicit task.

The present study demonstrated the specific brain regions for processing angry, fearful and happy prosody were left frontal pole / orbitofrontal cortex, right supramarginal gyrus, and left inferior frontal gyrus respectively, and the important role of right superior temporal gyrus and right supramarginal gyrus in emotional explicit task. These findings partially support the hierarchical model of emotional prosody and question the third level of the model.

Key words emotion; phonetic prosody; superior temporal gyrus; orbitofrontal cortex; inferior frontal gyrus; supramarginal gyrus